

Exercise 1: High-dimensional Gaussian Distributions

Consider a random vector $X = (X_1, \dots, X_p)^\top \sim \mathcal{N}(0, \mathbf{I})$, i.e., a multivariate normally distributed vector with mean vector zero and covariance matrix being the identity matrix of dimension $p \times p$. In this case, the coordinates X_1, \dots, X_p are i.i.d. each with distribution $\mathcal{N}(0, 1)$. Recall that the L_1 -norm of a vector $\mathbf{x} = (x_1, \dots, x_p)^\top \in \mathbb{R}^p$ is defined as

$$\|\mathbf{x}\|_1 = \sum_{i=1}^p |x_i|.$$

- (a) What is $\mathbb{E}[\|X\|_1]$? (Hint: $\mathbb{E}|X| = \sqrt{\frac{2}{\pi}}$.)
- (b) What is $\text{Var}[\|X\|_1]$? (Hint: $\text{Var}|X| = 1 - \frac{2}{\pi}$.)
- (c) Now let $X' = (X'_1, \dots, X'_p)^\top \sim \mathcal{N}(0, \mathbf{I})$ be another multivariate normally distributed vector with mean vector zero and covariance matrix being the identity matrix of dimension $p \times p$. Further, assume that X and X' are independent. What is $\mathbb{E}(\|X - X'\|_1)$?
- (d) What is $\text{Var}(\|X - X'\|_1)$?

(e) Let $\mathbf{x} = (x_1, \dots, x_p)^\top \in \mathbb{R}^p$ be some arbitrary deterministic vector. Compute $\mathbb{E}\langle X, \mathbf{x} \rangle$ and $\text{Var}(\langle X, \mathbf{x} \rangle)$.